



500.43870X00

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: H. KIBA, et al

Serial No.: 10/849,022

Filing Date: May 20, 2004

For: METHOD AND SYSTEM FOR DATA PROCESSING FOR
CONTROLLING A CACHE MEMORY

LETTER CLAIMING RIGHT OF PRIORITY

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

October 7, 2004

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55, applicants hereby claim
the right of priority based on:

Japanese Application No. 2004-132479
Filed: April 28, 2004

A Certified copy of said application document is attached hereto.

Acknowledgement thereof is respectfully requested.

Respectfully submitted,

Carl I. Brundidge
Registration No. 29,621
ANTONELLI, TERRY, STOUT & KRAUS, LLP

CIB/jdc
Enclosures
703/312-6600

日本国特許庁
JAPAN PATENT OFFICE

CERTIFIED COPY OF
PRIORITY DOCUMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出願年月日 2004年 4月28日
Date of Application:

出願番号 特願2004-132479
Application Number:

ST. 10/C]: [JP 2004-132479]

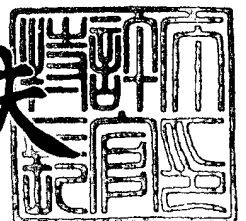
願 人 株式会社日立製作所
Applicant(s):

BEST AVAILABLE COPY

2004年 5月26日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



【書類名】 特許願
【整理番号】 K03013301A
【あて先】 特許庁長官殿
【国際特許分類】 G06F 12/08
【発明者】
 【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所
 ソフトウェア事業部内
 【氏名】 木場 弘康
【発明者】
 【住所又は居所】 神奈川県横浜市戸塚区戸塚町 5 0 3 0 番地 株式会社日立製作所
 ソフトウェア事業部内
 【氏名】 河村 信男
【特許出願人】
 【識別番号】 000005108
 【氏名又は名称】 株式会社 日立製作所
【代理人】
 【識別番号】 100075096
 【弁理士】
 【氏名又は名称】 作田 康夫
【選任した代理人】
 【識別番号】 100100310
 【弁理士】
 【氏名又は名称】 井上 学
【手数料の表示】
 【予納台帳番号】 013088
 【納付金額】 16,000円
【提出物件の目録】
 【物件名】 特許請求の範囲 1
 【物件名】 明細書 1
 【物件名】 図面 1
 【物件名】 要約書 1

【書類名】 特許請求の範囲**【請求項1】**

プログラムを実行する計算機と、前記プログラムの実行により送られたデータを格納するキャッシュメモリと該キャッシュメモリに格納されたデータを格納するためのディスク装置とを有する記憶装置とを備えたデータ処理システムにおけるキャッシュ制御方法において、

前記記憶装置は、前記プログラムから送られたデータの格納要求の入力に応じて前記送られたデータを前記キャッシュメモリに記憶し、前記プログラムから送られたフラッシュ要求の入力に応じて前記キャッシュメモリに格納されたデータを前記ディスク装置に格納することを特徴とするキャッシュ制御方法。

【請求項2】

前記フラッシュ要求は、前記プログラムの稼動によるトランザクション処理におけるチェックポイントのタイミングで前記プログラムから前記記憶装置へ送信することを特徴とする請求項1記載のキャッシュ制御方法。

【請求項3】

前記データ格納要求と前記フラッシュ要求は、前記キャッシュメモリの領域を特定する領域識別情報を含み、前記データ格納要求が入力されると該データ格納要求の領域識別情報で特定される領域に前記送られたデータを格納し、前記フラッシュ要求が入力されると該データ格納要求の領域識別情報で特定される領域に格納されたデータを前記ディスク装置へ格納することを特徴とする請求項1記載のキャッシュ制御方法。

【請求項4】

前記キャッシュメモリの領域は、データ更新が発生したか否かが管理され、前記フラッシュ要求が入力されると該データ格納要求の領域識別情報で特定される領域に格納されたデータが更新されたものを前記ディスク装置へ格納することを特徴とする請求項3記載のキャッシュ制御方法。

【請求項5】

前記キャッシュメモリの領域識別は、ボリューム識別情報およびセグメント識別情報であることを特徴とする請求項3記載のキャッシュ制御方法。

【請求項6】

プログラムを実行する計算機と、前記プログラムの実行により送られたデータを格納するキャッシュメモリと該キャッシュメモリに格納されたデータを格納するためのディスク装置とを有する記憶装置とを備えたデータ処理システムにおいて、

前記記憶装置は、前記プログラムから送られたデータの格納要求の入力に応じて前記送られたデータを前記キャッシュメモリに記憶する手段と、前記プログラムから送られたフラッシュ要求の入力に応じて前記キャッシュメモリに格納されたデータを前記ディスク装置に格納する手段とを備えたことを特徴とするデータ処理システム。

【請求項7】

前記フラッシュ要求は、前記プログラムの稼動によるトランザクション処理におけるチェックポイントのタイミングで前記プログラムから前記記憶装置へ送信することを特徴とする請求項6記載のデータ処理システム。

【請求項8】

前記データ格納要求と前記フラッシュ要求は、前記キャッシュメモリの領域を特定する領域識別情報を含み、前記データ格納要求が入力されると該データ格納要求の領域識別情報で特定される領域に前記送られたデータを格納し、前記フラッシュ要求が入力されると該データ格納要求の領域識別情報で特定される領域に格納されたデータを前記ディスク装置へ格納することを特徴とする請求項6記載のデータ処理システム。

【請求項9】

前記キャッシュメモリの領域は、データ更新が発生したか否かが管理され、前記フラッシュ要求が入力されると該データ格納要求の領域識別情報で特定される領域に格納されたデータが更新されたものを前記ディスク装置へ格納することを特徴とする請求項8記載のデ

ータ処理システム。

【請求項10】

前記キャッシュメモリの領域識別は、ボリューム識別情報およびセグメント識別情報であることを特徴とする請求項8記載のデータ処理システム。

【請求項11】

プログラムを実行する計算機と、前記プログラムの実行により送られたデータを格納するキャッシュメモリと該キャッシュメモリに格納されたデータを格納するためのディスク装置とを有する記憶装置とを備えたデータ処理システムを機能させるデータ処理プログラムにおいて、

前記記憶装置では、前記プログラムから送られたデータの格納要求の入力に応じて前記送られたデータを前記キャッシュメモリに記憶するステップとと、前記プログラムから送られたフラッシュ要求の入力に応じて前記キャッシュメモリに格納されたデータを前記ディスク装置に格納するステップとを有することを特徴とするデータ処理プログラム。

【書類名】 明細書**【発明の名称】** キャッシュ制御およびデータ処理システム並びにその処理プログラム**【技術分野】****【0001】**

本発明はキャッシュメモリを有する記憶装置のキャッシュフラッシュ処理を、記憶装置外部のタイミングで実施するキャッシュ制御技術に関する。

【背景技術】**【0002】**

特許文献1で述べられているように、電源が遮断されたタイミングで、ディスク記憶装置を停止するためにディスク記憶装置内のキャッシュメモリの内容を磁気ディスクに強制的にキャッシュをフラッシュするキャッシュ制御方式が開示されている。

【0003】**【特許文献1】** 特開平10-254780号**【発明の開示】****【発明が解決しようとする課題】****【0004】**

上記従来技術は、データ処理装置で稼動するプログラムから記憶装置に対してフラッシュ指示ができないという問題があった。

本発明の目的は、記憶装置外部からプログラムの処理のタイミングでキャッシュメモリから記録媒体への書き込みを実施することにある。

【課題を解決するための手段】**【0005】**

上記目的を達成するために、上位装置のアプリケーションプログラムの同期点処理において、記憶装置に付属するキャッシュへのフラッシュ指示をするようにして、アプリケーションの同期点処理とフラッシュ処理を連動させるようにしたものである。

【発明の効果】**【0006】**

本発明によれば、記憶装置外部からプログラムの処理のタイミングでキャッシュメモリから記録媒体への書き込みを実施することが可能となる。

【発明を実施するための最良の形態】**【0007】**

以下、本発明の実施例を図面に基づいて詳細に説明する。図1は本発明の一実施例である上位装置と記憶装置の構成を示すブロック図である。本実施例では、上位装置10と記憶装置11で構成されており、上位装置10ではアプリケーションプログラム100が稼動し、アプリケーションプログラム100には読出／書込処理部101と同期点処理部102が割り当てられている。書込／読出処理部101は記憶装置11に読出要求103や書込要求104を発行する。同期点処理部102は記憶装置11に書込要求105やフラッシュ指示106を発行する。記憶装置11にはキャッシュ制御部110とキャッシュメモリ120と記憶媒体130が割り当てられている。キャッシュ制御部110はキャッシュ管理テーブル117を参照してヒット判定113、キャッシュメモリメモリ120に対するキャッシュ読出113やキャッシュ書込114、記憶媒体130に対する媒体読出115や媒体書込116を行う。キャッシュメモリ120はボリューム毎の区画121に分かれる、更に細分化したLU単位の区画122（LU：論理ユニット）に分かれて、LU内はLBA（LBA：論理ブロックアドレス）が記憶されている。記憶媒体130は複数のボリューム131を有する。本実施例ではキャッシュメモリ120をボリューム単位の区画121やLU単位の区画122に分ける実施例を示したが、ボリューム単位の区画121やLU単位の区画122に分けなくてもよい。ここで説明した各処理部は、プログラムを一例として説明しているが、オブジェクトやプロセス、スレッド、ハードウェアなどで実現しても良い。

【0008】

データベース管理システムなどでは、障害に備え、定期的にDBの整合性を保証する為、稼動中にチェックポイントを取得する。チェックポイントは、システム障害時の再開始処理を行う起点となるデータベースの整合性が保証されたポイントとなる。主にチェックポイントは稼動中に出力したログブロック数がある一定の回数に達した時点に取得する場合が多い。チェックポイント処理では、その時点のデータベースバッファにおける更新が行われたデータベースブロックを記憶装置上に全て書き出す処理が行われる。アプリケーションプログラム100として、データベース管理システムが利用された場合、フラッシュ指示106として上記のチェックポイント処理が効力を発揮する。例えば、データベース管理システムで書き込んだデータを、フラッシュ指示106として上記のチェックポイントが発生したタイミングを用いてフラッシュ指示を実施することにより、チェックポイント毎にキャッシュメモリにおけるデータベース管理システムで書き込んだデータがディスクへフラッシュされることになる。このようにすることにより、電源断の発生やシステムを停止する場合に、ディスクへのフラッシュが行われているため、記憶媒体130を停止する時間を短縮することが可能となる。

【0009】

図2は同期点処理部102における同期点処理の流れを示す。アプリケーションプログラム100のメモリに保持され記憶装置11に未反映の更新データをwrite throughモードで書込要求104を出し(ステップ21)、記憶媒体130に未反映のキャッシュメモリ120上のページを記憶媒体130に反映するために記憶装置11にフラッシュ指示105を出し(ステップ22)、記憶装置11に当該同期点処理が終了したことを記録する同期点ジャーナルをwrite throughモードで記憶装置11に書込要求104を出す(ステップ23)。

【0010】

図3はキャッシュ制御部110によるフラッシュ指示処理の流れを示す。キャッシュ管理テーブル117のファイル管理部51を走査して最初のファイル501の情報を参照し(ステップ301)、フラッシュ指示105で指定されたボリュームとボリューム502が一致する又はボリュームの指定がない条件が成立しなければ次のファイル501の走査に戻り、条件が成立すれば次の処理に進み(ステップ302)、フラッシュ指示105で指定されたLUとLU503が一致する又はLUの指定がない条件が成立しなければ次のファイル501の走査に戻り、条件が成立すれば次の処理に進み(ステップ303)、モード505がwrite afterに一致しなければ記憶媒体130に反映されているので次のファイル501の走査に戻り、モード505がwrite afterに一致すれば次の処理に進み(ステップ304)、キャッシュメモリポインタ506がキャッシュ管理エントリ52を指していなければ次のファイル501の走査に戻り、キャッシュメモリポインタ506がキャッシュ管理エントリ52を指していれば次の処理に進む(ステップ305)。

【0011】

次にキャッシュ管理エントリ53を参照し(ステップ306)、当該キャッシュ管理エントリ53の指すページ512が記憶媒体130に反映済であれば次のキャッシュ管理エントリ52の走査に戻り、記憶媒体130へ未反映であれば次の処理に進み(ステップ307)、当該キャッシュ管理エントリ53の指すページ512を記憶媒体130に媒体書込116を行って、キャッシュ管理エントリ53を反映済に変更し(ステップ308)、ネクストポインタ511があれば次のキャッシュ管理エントリ52の走査に戻り、ネクストポインタ511がなければ次の処理に進み(ステップ309)、ファイル管理部51の全てのファイルを走査していなければ次のファイル501の走査に戻り、全てのファイルを走査していればフラッシュ指示処理を終了する(ステップ310)。

【0012】

図4はキャッシュ制御部110による書込要求処理の流れを示す。キャッシュ管理テーブル117のファイル管理部51を走査して書込要求104で指定されたファイルと一致するファイル501を検索してキャッシュポインタ506を取得して(ステップ401)

、書き込みデータと突き合わせて対応するキャッシュ管理エントリ53を検索し（ステップ402）、書き込みデータに対応するキャッシュ管理エントリ53がヒットすれば次の処理に進み、書き込みデータに対応するキャッシュ管理エントリ53がヒットしなければ（ステップ403）、新規キャッシュ管理エントリ53及びキャッシュメモリページの確保とキャッシュ管理エントリ53のリストメンテナンスを行い（ステップ404）、書込要求104時に指定されたモードがwrite afterであればステップ406の処理に分岐し（ステップ405）、書き込みデータをキャッシュメモリ120にキャッシュ書込114して（ステップ406）、キャッシュ管理エントリ53を未反映512に変更し（ステップ408）、ステップ410の処理に進む。一方、書込要求104時に指定されたモードがwrite throughであればステップ407の処理に分岐し（ステップ405）、書き込みデータをキャッシュメモリ120にキャッシュ書込114すると共に記憶媒体130に媒体書込116を行い（ステップ407）、キャッシュ管理エントリ53を反映済512に変更し（ステップ409）、ステップ410の処理に進む。書き込みデータの全てについてキャッシュ管理エントリ53と突合せ検索が終っていなければ突合せ検索に戻り、終っていれば次の処理に進み（ステップ410）、書込要求104時に指定されたモードをモード505に設定し（ステップ411）、書込要求処理を終了する。

【0013】

図5はキャッシュ制御部110におけるキャッシュ管理テーブル117のデータ構造図を示す。キャッシュ管理テーブル117はファイル管理部51とキャッシュ管理部52からなる。ファイル管理部51において、ファイル501はファイル名を保持し、ボリューム502はファイルの格納ボリュームを示し、LU503は格納LUを示し、LBA504は格納LBAを示し、モード505はファイル書込要求時に記憶媒体130とキャッシュメモリ120の両方に書き込み動作を行うwrite throughモードと書込要求時はキャッシュメモリ120のみに書き込み、フラッシュ時に記憶媒体140に書き込み動作を行うwrite afterモードがある。キャッシュポインタ506は、当該ファイルで使用しているキャッシュ管理エントリ53を指すリストの先頭ポインタを保持する。キャッシュ管理部52は対応するキャッシュページの情報を持つ複数のキャッシュ管理エントリ53からなる。ネクストポインタ511は当該ファイル501の次のキャッシュ管理エントリ53を指す。ページ512はキャッシュメモリ120内のページを示す。LBA513は記憶媒体130の格納位置を示す。反映済又は未反映514は当該キャッシュメモリページが記憶媒体130に反映された状態（内容が一致する）であるか、反映されていない状態（内容が不一致）であるかを保持する。

【0014】

以上によれば、記憶装置のキャッシュメモリの内容が失われる障害発生してもアプリケーションプログラムの同期点までの更新内容が記録媒体に書き込まれていることを保証できるため、障害後のアプリケーションプログラムの回復処理では回復開始地点を正確に決めることができる。

【0015】

また、アプリケーションプログラムの制御により逐次キャッシュフラッシュが行えるため、大量のキャッシュメモリを有する記憶装置における電源停止時にアプリケーションの制御に因らずにキャッシュに蓄積された大量の更新データのフラッシュ処理より、停止処理に時間を要する問題を回避することができる。

【図面の簡単な説明】

【0016】

【図1】本発明の一実施例である上位装置と記憶装置の構成を示すブロック図である。

【図2】本発明の一実施例である上位装置のアプリケーションの同期点処理を示すフローチャートである。

【図3】本発明の一実施例による記憶装置のフラッシュ指示処理を示すフローチャートである。

トである。

【図 4】 本発明の一実施例による記憶装置の書込要求処理を示すフローチャートである。

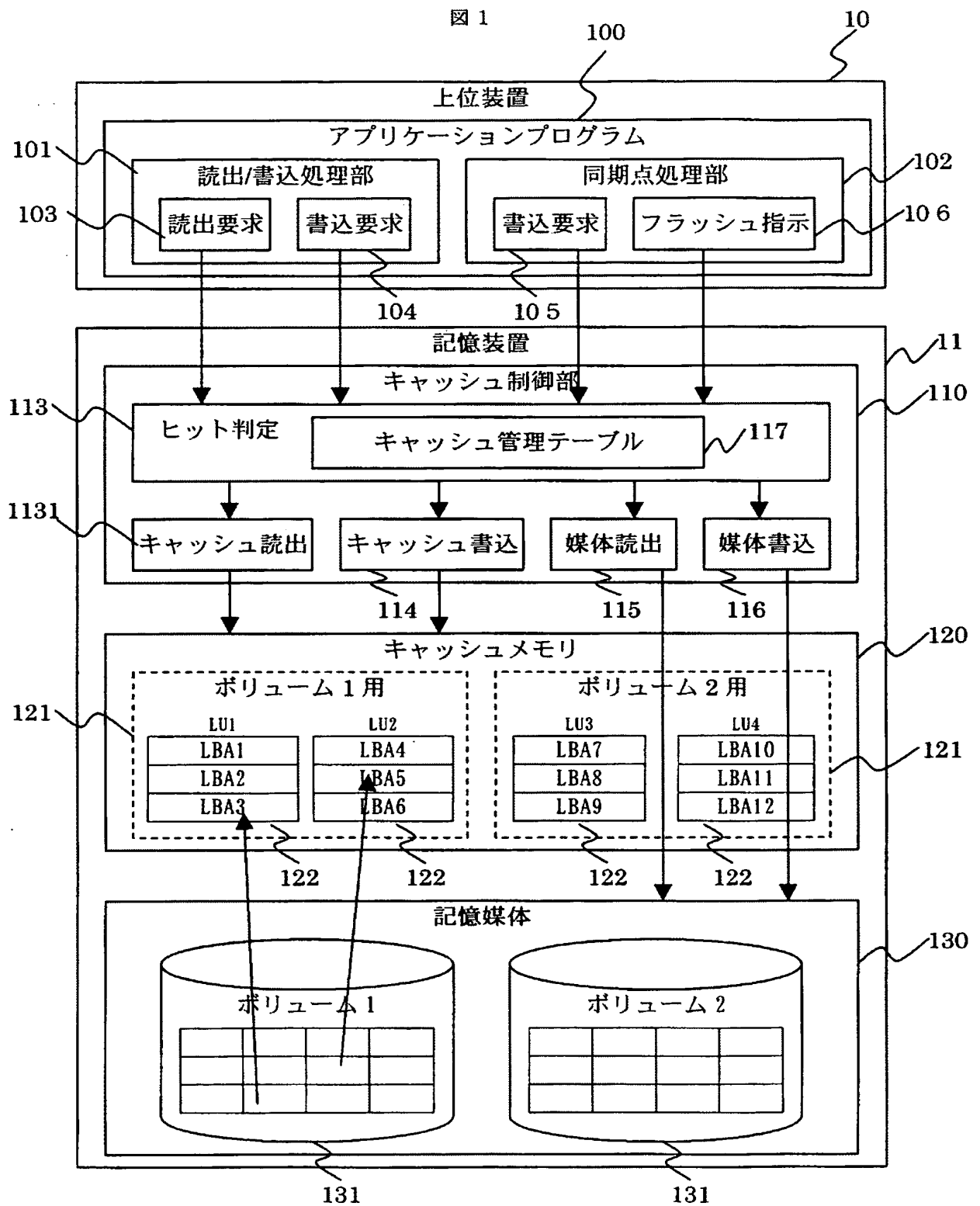
【図 5】 本発明の一実施例によるキャッシュ管理テーブルのデータ構造図である。

【符号の説明】

【 0 0 1 7 】

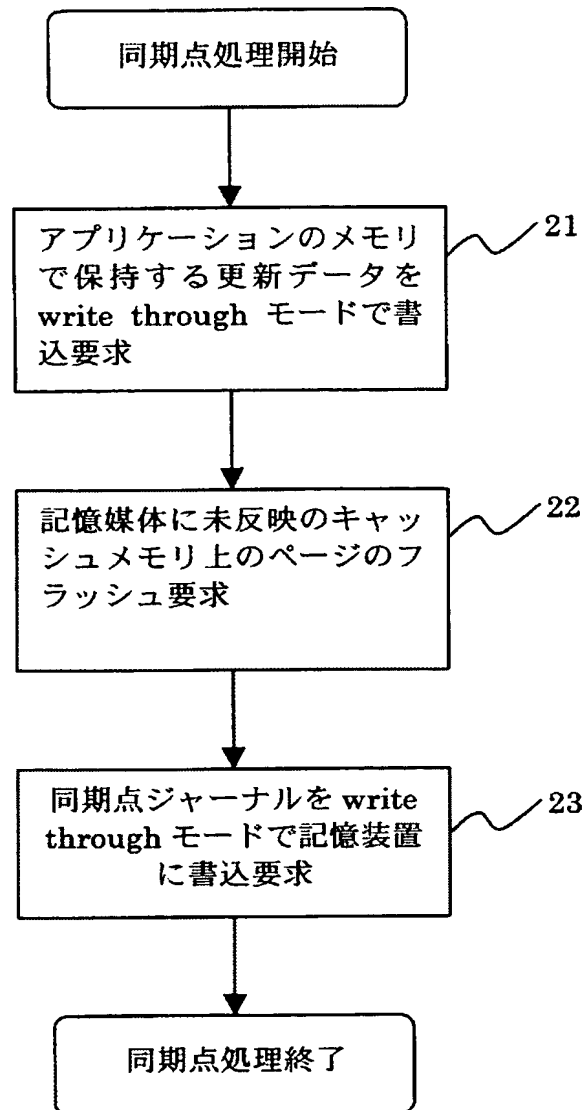
- 1 0 上位装置
- 1 0 0 アプリケーションプログラム
- 1 0 1 読出／書込処理部
- 1 0 2 同期点処理部
- 1 1 記憶装置
- 1 1 0 キャッシュ制御部
- 1 2 0 キャッシュメモリ
- 1 3 0 記憶媒体

【書類名】 図面
【図 1】

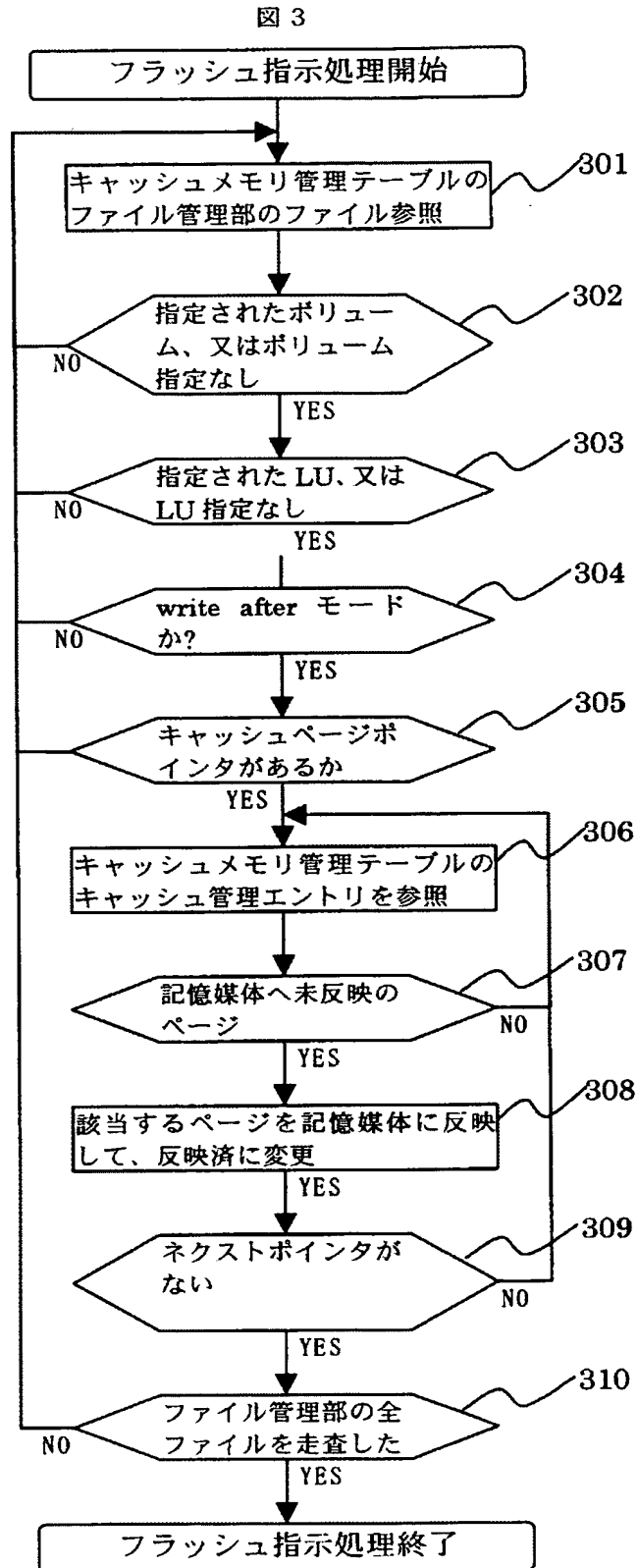


【図2】

図2

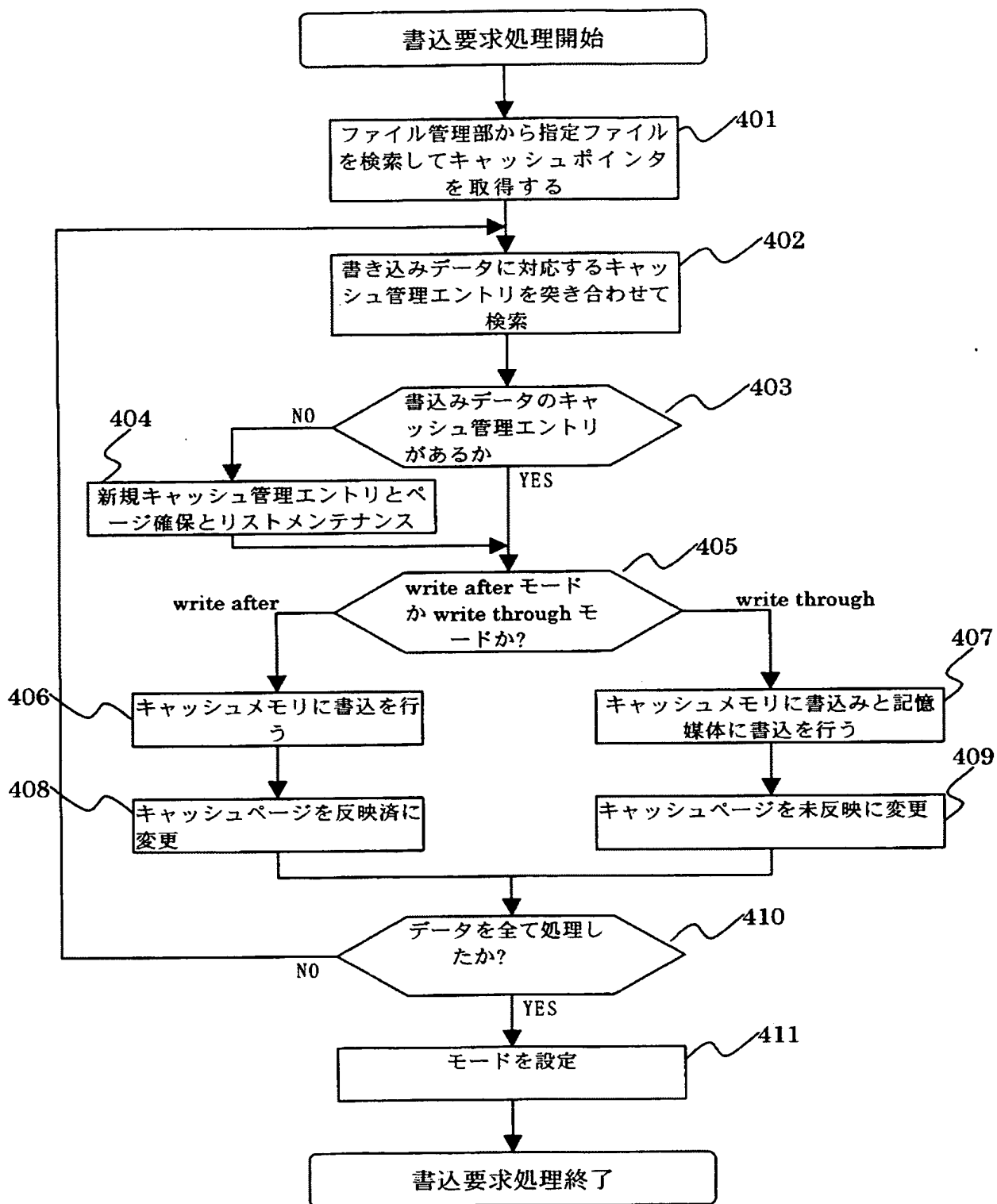


【図 3】



【図4】

図4



【書類名】 要約書**【要約】****【課題】**

本発明はキャッシュメモリを有する記憶装置のキャッシュフラッシュ処理と、上位装置の同期点処理を行うプログラムに関するものであり、従来は、上位装置のプログラムは記憶装置に対してフラッシュ指示ができないという問題があった。

【解決手段】

プログラムを実行する計算機と、前記プログラムの実行により送られたデータを格納するキャッシュメモリと該キャッシュメモリに格納されたデータを格納するためのディスク装置とを有する記憶装置とを備えたデータ処理システムを用いる。前記記憶装置は、前記プログラムから送られたデータの格納要求の入力に応じて前記送られたデータを前記キャッシュメモリに記憶する。次に、前記プログラムから送られたフラッシュ要求の入力に応じて前記キャッシュメモリに格納されたデータを前記ディスク装置に格納する。

【選択図】 図1

認定・付加情報

特許出願の番号	特願 2 0 0 4 - 1 3 2 4 7 9
受付番号	5 0 4 0 0 7 2 9 2 0 8
書類名	特許願
担当官	第七担当上席 0 0 9 6
作成日	平成 1 6 年 4 月 3 0 日

< 認定情報・付加情報 >

【提出日】	平成 1 6 年 4 月 2 8 日
-------	--------------------

特願 2 0 0 4 - 1 3 2 4 7 9

出 願 人 履 歴 情 報

識別番号

[0 0 0 0 0 5 1 0 8]

1. 変更年月日

1 9 9 0 年 8 月 3 1 日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台 4 丁目 6 番地

氏 名

株式会社日立製作所

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☒ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.